Quasi-Local Mass in General Relativity

Shing-Tung Yau

It is a great honor for me to give this talk in honor of my mentor and my old friend Is Singer. We have known each other since 1973. He had a deep influence in my mathematical thinking and my career. When I first met him at a big conference in geometry in Stanford, my advisor Chern told me that Is wanted to see me. I was somewhat nervous as Is is a big name in geometry. To my surprise, Is just smiled and said that he just wanted to get to know me. But he did something for me then that neither of us realized at the time. It was his proposal to invite Robert Geroch to explain the mathematics of general relativity to geometers during that conference. Geroch explained the positive mass conjecture for a special case in very simple terms. It was fascinating as nobody knew how to approach this problem, despite its elegance.

I kept this problem in mind until 1978, when I visited Berkeley for one year, again at the invitation of Chern and of Is. In that year, I solved the first major part of the positive conjecture with Richard Schoen. Before this visit, Is invited me to visit MIT for a month during 1976, after I solved the Calabi conjecture. During that period, Is was very busy with a family issue. Nonetheless, we had dinner and he described to me what he heard in Stony Brook when he visited the physicists there. The main question was concerning self-dual instantons. Although I did not work on it immediately, it is such an interesting subject, that I finally worked on the higher dimensional generalization of it known today as the Hermitian Yang-Mills connections.

In 1987, Is thought that it was important for mathematicians working on physics related issues to have the ability to obtain funding to hire postdoctoral fellows. He was able to work this out with the Department of Energy. And I have benefitted much by his efforts on this. I was able to hire postdoctoral fellows who earned physics degrees. Having them around have been really useful for my career.

Talk given at the Perspectives in Mathematics and Physics Conference, in celebration of Professor I.M. Singer’s 85th birthday.

©2011 International Press
Is and I have worked together on numerous occasions, including the time that we initiated the mirror symmetry workshop in MSRI. There, we were able to attract the interest of a whole generation of mathematicians and get them excited about mirror symmetry. The mathematics inspired by physics is truly exciting and inspiring. It was Is Singer who helped me to direct my research in this beautiful direction. It is almost forty years since I met Is. I am truly grateful to his leadership in this field.

Let me now begin on my recent work with Muo-Tao Wang on general relativity. The ideas went back to the work with Schoen on positive mass conjecture. In general relativity, Einstein’s equation is obtained by taking the variation of

$$\frac{1}{16\pi} \int R + \int L$$

where $R$ is the scalar curvature of the spacetime and $L$ is the Lagrangian of matter coupled to gravity. The gravitational interaction is described by means of a spacetime Lorentzian metric $g_{ij}$ which has indefinite signature $(-,+,+,+)$. For instance, the metric of the Minkowski spacetime $\mathbb{R}^{3,1}$ which is the vacuum with zero matter is

$$ds^2 = g_{ij} dx^i dx^j = -dt^2 + (dx)^2 + (dy)^2 + (dz)^2.$$  

The variational equation has the form

$$R_{ij} - \frac{1}{2} R g_{ij} = T_{ij}$$

where $R_{ij}$ is the Ricci tensor, and $T_{ij}$ is called the matter energy-momentum tensor.

In classical relativity, the matter tensor satisfies the weak energy condition

$$\sum T_{ij} l^i l^j \geq 0$$

for any four-vector $l^i$ that is time-like

$$\sum g_{ij} l^i l^j < 0.$$

There are two very important solutions of the Einstein equation that are used extensively in general relativity. The Schwarzschild metric

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2)$$

and the Kerr metric

$$ds^2 = -\frac{\Delta}{U} (dt - a \sin^2 \theta d\phi)^2 + U \left(\frac{dr^2}{\Delta} + d\theta^2\right) + \frac{\sin^2 \theta}{U} (a dt - (r^2 + a^2) d\phi^2)^2$$
where

\[ U = r^2 + a^2 \cos^2 \theta \]
\[ \Delta = r^2 - 2Mr + a^2 \]
\[-\infty < t < \infty, M + \sqrt{M^2 - a^2} < r < \infty, \]
\[ 0 < \theta < \pi, 0 < \phi < 2\pi. \]

Both solve the Einstein equation with \( T_{ij} = 0 \) (vacuum). The first one is spherically symmetric and the second one is axially symmetric. They respectively describe static and stationary black holes. The constants \( M \) and \( a \) are the mass and angular momentum of the black hole. The famous black hole uniqueness theorem says that stationary black holes with no matter are exactly described by the Kerr metric.

Both metrics have a null hypersurface at

\[ r = M + \sqrt{M^2 - a^2} \]

which is the event horizon of the black hole. This is the spacetime boundary of the black hole where any event occurring inside can not be detected by an outside observer. Note that the vector field \( \frac{\partial}{\partial t} \) is a Killing vector field as it preserves the metric.

The Killing field \( \frac{\partial}{\partial t} \) is time-like (i.e. \( g(\frac{\partial}{\partial t}, \frac{\partial}{\partial t}) < 0 \)) when

\[ r^2 - 2Mr + a^2 \cos^2 \theta > 0 \]

but space-like (i.e. \( g(\frac{\partial}{\partial t}, \frac{\partial}{\partial t}) > 0 \)) when

\[ r^2 - 2Mr + a^2 \cos^2 \theta < 0. \]

This last region is called the ergosphere. It is a bounded region outside the event horizon except at \( \theta = 0 \) and \( \pi \).

We can consider the dynamics of a scalar field \( \Phi(t, r, \theta, \phi) \) in the Kerr spacetime. Its propagation is described by a scalar wave equation. Since the Kerr spacetime has a Killing vector field \( \frac{\partial}{\partial t} \), the Lagrangian \( \int |\nabla \Phi|^2 \) associated to the wave equation defines a local energy density. It has the form

\[ E = \left( \frac{(r^2 + a^2)^2}{\Delta} - a^2 \sin^2 \theta \right) |\partial_t \Phi|^2 + \Delta |\partial_r \Phi| + \sin^2 \theta |\partial_{\cos \theta} \Phi|^2 \]
\[ + \left( \frac{1}{\sin^2 \theta} - \frac{a^2}{\Delta} \right) |\partial_{\phi} \Phi|^2 \]

This density is positive except within the ergosphere, where it is negative. In the Kerr geometry, no first-order or higher-order positive conserved energy density exists for the scalar wave equation as was observed by Finster, Kamran, Smoller and Yau [7]. The major classical method to study wave equation is called the energy method. One exploits the conservation of total energy and Sobolev type inequality to control local behavior in terms of the
total energy. However, the energy method for the scalar wave equation breaks down due to the negativity of the energy density within the ergosphere unless the angular momentum is small relative to the mass. (In this case, Dafermos and Rodnianski [5] proved the solution is bounded in $t$ if the initial data has compact support outside the event horizon.)

An important question for the wave equation is whether the wave will decay in time if initially it does not spread out. The decay of the solution of the scalar wave equation is a special case of the decay of solutions to the Teukolsky equation which describes the linear stability of the Kerr black hole. In this connection, we quote Frolov and Novikov [9]:

"Linear stability of the Kerr black hole is one of the few truly outstanding problems that remain in the field of black hole under gravitational wave perturbations."

The problem of linear stability of Kerr is still open and it is not even clear whether linear stability of the Kerr holds true if the angular momentum is not small relative to the mass.

A remarkable property of the Kerr metric is that the Teukolsky equation is separable by the ansatz

$$e^{-i\omega t - ik\phi} R(r) \Theta(\theta)$$

Chandrasekhar called this property of Kerr geometry “having the aura of the miraculous” [3]. Making use of this, in 2001, Finster, Kamran, Smoller and Yau [6] proved that the propagation of waves described by a Dirac equation in Kerr space decays in time like $t^{-5/6}$. We also proved that for the scalar wave equation, the wave with fixed angular momentum mode $k$ also decays.

In principle, we can sum up the modes to conclude the decay of the scalar wave equation. This can be done for the Schwarzschild geometry. However, the negativity of the energy density for the scalar wave in the ergosphere of the Kerr geometry causes problems. Indeed, the ergosphere has many strange properties including the energy extraction process proposed by Penrose [18]. The wave analogue was proposed by Zeldovich [29] and Starobinsky [23]. It is called superradiance. A complete rigorous treatment for the latter case was finally achieved recently by Finster, Kamran, Smoller and Yau [8].

We have seen now that in the ergosphere, the energy density of the scalar wave can be negative and has give peculiar properties. The problem comes from the fact that the gravitational field itself must have energy. After all, the potential energy of a pair of gravitating particles depends on their separation distance. Hence the total energy depends on the gravitational field configuration.

In special relativity, the tensor $T_{ij}$ can be used to define the energy-momentum vector for a domain in space in the following manner: Take any Killing vector $v^j$ which preserves the metric.

$$J_i = T_{ij}v^j$$
defines a divergence free current. The Hodge dual of $J$ is a closed three form. When we integrate this three form over a compact space-like hypersurface, we obtain the energy-momentum vector $P$ associated to this hypersurface.

However, since $T_{ij}$ accounts only for matter, the energy-momentum vector defined in this way does not account for the gravitational energy. A more serious problem is that a general spacetime does not admit any Killing field. If the spacetime admits a time-like Killing field, Noether’s theorem applied to the Lagrangian provides a current associated to this Killing field. This current can be used to define a mass called the Komar mass. Let us now recall Noether’s theorem:

Let $L$ be the Lagrangian which can be considered as a function defined on the tangent bundle of a manifold $M$. Suppose we have a one parameter family of diffeomorphisms $h_t : M \rightarrow M$ so that

$$L((h_t)_* v) = L(v)$$

Then the Euler-Lagrange equation associated to $L$ admits a first integral $I : TM \rightarrow \mathbb{R}$ given by

$$I(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}} \frac{db_t(q)}{dt} \bigg|_{t=0}.$$ 

For most mechanical systems, the Lagrangian is invariant under translation of time and space and the resulting conserved quantity is a four-vector defining energy and linear momentum. However, in general relativity, most spacetimes do not admit translational symmetry, so it is difficult to define energy-momentum using Noether’s theorem. However, one can define a total energy-momentum vector for an isolated physical system if there is asymptotic translational symmetry.

When the spacetime is asymptotically flat, there is a space-like hypersurface which outside a compact set is diffeomorphic to $\mathbb{R}^3$ minus a ball and the metric $g_{ij}$ has the form

$$g_{ij} = \delta_{ij} + \mathcal{O}\left(\frac{1}{r}\right),$$

and the second fundamental form $p_{ij}$ of this hypersurface satisfies:

$$p_{ij} = \mathcal{O}\left(\frac{1}{r^2}\right).$$

We expect that for such an asymptotically-flat spacetime, there is Lorentzian symmetry at infinity, and the approximate symmetry at infinity can be used to define the total mass (or energy)

$$M_{ADM} = \frac{1}{16\pi} \lim_{r \to \infty} \int_{S^2(r)} (\partial_i g_{ij} - \partial_j \text{tr} g) v^j.$$
and the total linear momentum

\[ P_i = \lim_{r \to \infty} \frac{1}{8\pi} \int_{S^2(r)} (p_{ij} - p_{kk} \delta_{ij}) v^j. \]

The four-vector \( \left( \frac{M}{P_i} \right) \) is called the ADM (Arnowitt-Deser-Misner) energy-momentum four-vector. It was proved to be a time-like vector except for Minkowski spacetime which of course is trivially null.

The total energy in general relativity cannot be obtained by integrating any local density along a hypersurface. The reason is that the density would depend on the first order differentiation of the metric \( g_{ij} \). But there is a coordinate system where such quantities are zero at that point. The equivalence principle says that physical quantities should be independent of the choice of the coordinate system. Nonetheless, one can still ask whether it is still possible to have a quasi-local mass, where a total energy-momentum four-vector is assigned to any space-like sphere bounding a compact portion of a space-like hypersurface.

In 1982, Penrose [20] listed the search for a definition of such quasi-local mass as his number one problem in classical general relativity besides his famous question of cosmic censorship.

There are many reasons to search for such a concept. Many important statements in general relativity make sense only with the presence of a good definition of quasi-local mass. For example, it allows us to talk about the binding energy of two bodies rotating around each other. More importantly, a good definition of quasi-local mass should help us to control the dynamics of the gravitational field. Hopefully, this may be used to generalize the energy method in hyperbolic equations where difficulties were encountered even in the study of linearized stability of the Kerr metric.

There were many attempts, including approaches given by Penrose [19], Hawking [11], and Brown-York [2] to give the definition of quasi-local mass. We shall use an approach which seems to be most promising.

Recall that for a Lorentzian manifold \( M \) with boundary \( \partial M \), the action in general relativity should be

\[ I(g, \Phi) = \int_M \left( \frac{1}{16\pi} R + L(g, \Phi) \right) + \frac{1}{8\pi} \int_{\partial M} K \]

where \( K \) is the trace of the second fundamental form of \( \partial M \). The last term is needed to give rise to the right variational equation if we fix the metric and the matter field on the boundary.

If we demand that a certain background \((g_0, \Phi_0)\) is a static solution to the field equation, we replace \( I \) by

\[ I(g, \Phi) - I(g_0, \Phi_0). \]
Hence, for flat spacetime background with $g_0$ being flat and $\Phi_0 = 0$, we use
\[
\int_M \left( \frac{1}{16\pi} R + L(g, \Phi) \right) + \frac{1}{8\pi} \int_{\partial M} (K - K_0) .
\]

Suppose we take a family of space-like surface $\Sigma_t$ generated by a time-like vector field $t$ such that $t^\mu \nabla_\mu t = 1$. We can write
\[
t^\mu = N n^\mu + N^\mu
\]
where $n^\mu$ is the normal to $\Sigma_t$, $N$ is called the lapse function, $N^\mu$ is called the shift vector. In this notation, we find
\[
I(g, \Phi) = \int N dt \left[ \frac{1}{16\pi} \int_{\Sigma_t} \left( R + p_{\mu\nu} p^{\mu\nu} - p^2 + 16\pi L \right) + \frac{1}{8\pi} \int_{S_t^2} 2K \right],
\]
where $p_{\mu\nu}$ is the second fundamental form of $\Sigma_t$ and $p$ is its trace, $2K$ is the mean curvature of $\partial \Sigma_t = S_t$.

If one introduces the canonical momenta $k^{\mu\nu}$, $k$ conjugate to $3^\mu g_{\mu\nu}, \Phi$, we can rewrite the action to be
\[
\int dt \int_{\Sigma_t} \left( k^{\mu\nu} \dot{g}_{\mu\nu} + k \dot{\Phi} - N \mathcal{H} - N^\mu \mathcal{H}_\mu \right) + \frac{1}{8\pi} \int_{S_t} \left( N^2 2K - N^\mu p_{\mu\nu} r^\nu \right),
\]
where $\mathcal{H}$ is the Hamiltonian constraint
\[
T_{00} - \frac{1}{2} \left( R - p_{\mu\nu} p^{\mu\nu} + p^2 \right),
\]
and $\mathcal{H}_\mu$ is the momentum constraint
\[
T_{0\mu} - p_{\mu\nu,\nu} + p_\mu.
\]
Note that $\mathcal{H} = 0$ and $\mathcal{H}_\mu = 0$ when the equation of motion is satisfied. We will let $r^\nu$ to be the spacelike unit normal of $S_t$ which is tangent to $\Sigma_t$.

The Hamiltonian is then derived to be
\[
H = \int_{\Sigma_t} \left( N \mathcal{H} + N^\mu \mathcal{H}_\mu \right) - \frac{1}{8\pi} \int_{S_t} \left( N^2 2K - N^\mu p_{\mu\nu} r^\nu \right).
\]
If we take the background so that $p_{\mu \nu} = 0$, we see that the Hamiltonian relative to the background is given by
\[
\int_{\Sigma_t} (N \mathcal{H} + N^\mu \mathcal{H}_\mu) - \frac{1}{8\pi} \int_{S_t} (N(2K - 2K_0) - N^\mu p_{\mu \nu} r^{\nu})
\]
Hence associated to each time-like vector field $t$, we have the physical Hamiltonian
\[
-\frac{1}{8\pi} \int_{S_t} (N(2K - 2K_0) - N^\mu p_{\mu \nu} r^{\nu})
\]
This expression was derived by Brown-York [2] and Hawking-Horowitz [12]. They proposed to simply choose $N = 1$, $N^\mu = 0$ for the definition of quasi-local mass. In general, the definition does not give positivity except in the time symmetric case ($p_{\mu \nu} = 0$) which was proved by Shi-Tam [17].

The definition of Brown-York is gauge dependent. Liu-Yau [14, 15] defined a gauge independent mass to be
\[
-\frac{1}{8\pi} \int_S \left( \sqrt{(2K)^2 - (tr_{SP})^2} - 2K_0 \right)
\]
and proved that it is positive whenever the mean curvature vector of $S$ is space-like and the Gauss curvature is positive. The proof combined arguments of Schoen-Yau [21, 22] and Witten [28]). We needed to handle metrics where the mean curvature may jump along the boundary. The discontinuity of the Dirac spinor required nontrivial analysis. Note that $\sqrt{(2K)^2 - (tr_{SP})^2}$ is the Lorentian norm of the mean curvature vector
\[
H = -2Kr^{\nu} + (tr_{SP})n^{\nu}.
\]

Let me now describe the work that I did with Mu-Tao Wang [24, 25, 26, 27].

Given a surface $S$, we assume that its mean curvature is positive. We embed $S$ isometrically into $\mathbb{R}^{3,1}$. Given any constant unit future time-like vector $w$ (observer) in $\mathbb{R}^{3,1}$, we can define a future directed time-like vector field $\bar{w}$ along $S$ by requiring
\[
\langle H_0, w \rangle = \langle H, \bar{w} \rangle
\]
where $H_0$ is the mean curvature vector of $S$ in $\mathbb{R}^{3,1}$ and $H$ is the mean curvature vector of $S$ in spacetime. This means that the expansion of the body in $\mathbb{R}^{3,1}$ relative to the observed $w$ is the same as the expansion of the body relative to $\bar{w}$.

Note that given any surface $S$ in $\mathbb{R}^{3,1}$ and a constant future time-like unit vector $w^\nu$, there exists a canonical gauge $n^\mu$ (future timelike unit normal along $S$) such that
\[
\int_S N^2K_0 + N^\mu (p_0)_{\mu \nu} r^{\nu}
\]
is equal to the total mean curvature of $\hat{S}$, the projection of $S$ onto the orthogonal complement of $w^\mu$, $w^\mu = N n^\mu + N^\mu$, $r^\mu$ is the spacelike unit normal orthogonal to $n^\mu$, $p_0$ is the second fundamental form calculated by the three surface defined by $S$ and $r^\mu$.

From the matching condition and the correspondence $(w^\mu, n^\nu) \rightarrow (\overline{w}^\mu, \overline{n}^\nu)$, we can define a similar quantity from our data in spacetime

$$\int_S N^2 \overline{K} + N^\mu (\overline{p})_{\mu\nu} \overline{n}^\nu.$$ 

Making use of the work of Liu-Yau [14, 15], Wang and I proved that

$$8\pi E(w) = \int_S N^2 \overline{K} + N^\mu (\overline{p})_{\mu\nu} \overline{n}^\nu - \int_S N^2 K_0 + N^\mu (p_0)_{\mu\nu} r^\nu$$

is non-negative [25, 26].
We define a quasi-local mass to be
\[
\inf E(w)
\]
where the infimum is taken among all isometric embeddings into \(\mathbb{R}^{3,1}\) and timelike unit constant vector \(w \in \mathbb{R}^{3,1}\).

In summary, given a closed spacelike 2-surface in spacetime whose mean curvature vector is space-like, we associate an energy-momentum four-vector to it that depends only on the first fundamental form, the mean curvature vector and the connection of the normal bundle such that

1. It is Lorentzian invariant.
2. It is trivial for surfaces sitting in Minkowski spacetime and future time-like for surfaces in spacetime which satisfies the local energy condition.

We shall compare our definition with the classical definitions. First of all, let us look at some special cases.

Spherical symmetric spacetime are foliated by the orbits of \(SU(2)\). We can define a function on the spacetime by associating to its orbit the area \(4\pi r^2\). The mean curvature vector of the orbit is
\[
-\frac{2}{r} \nabla r
\]
where \(\nabla\) is with respect to the quotient Lorentzian \((1,1)\) metric. If this vector is space-like, the quasi-local mass of this orbit sphere is
\[
M = r (1 - |\nabla r|).
\]

Note that in 1964, Misner and Sharp [16] defined a mass
\[
m = \frac{r}{2} \left(1 - |\nabla r|^2\right)
\]
which is the same as the Hawking mass [11]
\[
\sqrt{\frac{A}{16\pi}} \left(1 - \frac{1}{16\pi} \int_S |H|^2\right).
\]
The relation with our mass is
\[
m = M - \frac{M^2}{2r}.
\]
From the formula of quasi-local mass, which we proved to be positive, we derived a corollary that the mass \(m\) (Hawking mass) is also positive. (This was proved by Christodoulou and Yau [4] under extra assumptions.)

Note that
\[
\frac{1}{2} M \leq m \leq M
\]
where on the apparent horizon $M = 2m$, and at space-like infinity $M = m$. Hence our quasi-local mass is equivalent to the standard definition in the case of spherically symmetric spacetime.

In the spatial direction, the Hawking mass is monotonically increasing along the inverse mean curvature flow \[10\] and this is important in Huisken-Ilmanen's work \[13\]. The quasi-local mass is not monotonically increasing in this sense. However, the spherical symmetric case indicates that such property may still hold, up to a constant depending on the initial surface.

In the future time-like or null direction, the quasi-local mass is expected to decrease up to a constant depending on the initial surface if we choose the equation of motion for the 2-surfaces carefully. In the case when $p_{ij} \equiv 0$, there is also a definition of the quasi-local mass by Bartnik \[1\] which is obtained by minimizing the ADM mass among all asymptotically flat extension of the data which does not contain an apparent horizon and which extends the original data.

Our definition of the quasi-local mass also satisfies the following important properties:

3. When we consider a sequence of spheres on an asymptotically flat space-like hypersurface, in the limit, the quasi-local mass (energy-momentum) is the same as the well-understood ADM mass (energy-momentum).
4. When we take the limit along a null cone, we obtain the Bondi mass.

These properties of the quasi-local mass is likely to characterize the definition of quasi-local mass, i.e. any quasi-local mass that satisfies all these properties may be equivalent to the one that we have defined.

Strictly speaking, we associate each closed surface not a four-vector, but a function defined on the light cone of the Minkowski spacetime. Note that if this function is linear, the function can be identified as a four-vector. It is a remarkable fact that for the sequence of spheres converging to spatial infinity, this function becomes linear, and the four-vector is defined and is the ADM four-vector that is commonly used in asymptotically flat spacetime.

In any case, we are still in the process of deriving more properties of the quasi-local mass that we just introduced. We hope it can be used in applications for astronomy and physics.

References


DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, MA 02138, USA